MONOCAMERAL VISUAL RECOGNITION OF MARCUS HAND POSTURES FOR PERSONAL ROBOTIC ASSISTANTS

M. Gonzalo-Tasis[1], C. Laschi[2], J. Finat[3] and P. Dario[2]
[1]Eca-Simm Group, Dept. of Computer Science, University of Valladolid, Spain
[2]ARTS Lab, Scuola Superiore Sant' Anna, Pisa, Italy
[3]Dept. of Algebra, Geometry and Topology, University of Valladolid, Spain

Abstract- **The postures recognition is the first step for the gestures tracking of an artificial or a natural hand. In this article, we show a visuo-motor tracking of mobile hand configurations, which is based on symbolic representations able of supporting the biomechanical and perceptual information relative to evolving postures. After recognition, we have a virtual skeleton to identify simulated artificial hands. Such postures identification is adapted to an artificial Marcus hand developed as a human prosthesis. Nevertheless, the complex mechatronic device, this symbolic representation allows the visual identification and tracking of hand points of interest, such as tactile sensors and finger joints. In this way, a feedback for the perception-action cycle is obtained to improve the man-machine interaction in Personal Robotics, with special regard to the assistance of disabled people.**

I. INTRODUCTION

The visual servoing problem requires a coordination between visual perception and motor tasks; involving mainly to mechatronic devices of an artificial hand. There are two main approaches: model-based and appearance-based; they would correspond to the classical old distinction between top-down and bottom-up algorithms. The 3D approach is obviously richer than the planar one, and avoids some problems related to the self-occlusion. However, the 3D reconstruction of articulated models is considerably harder, and it depends strongly on the existence of good models and accurate parameter estimation. The choice depends on the goal and the real-time requirements.

The identification, tracking, errors correction and optimised learning of postures can be focused to solve non-verbal communication in human-computer interaction [8] or to solve positioning, grasping and handling the objects in Robotics. Non-verbal communication depends on people and cultures, and there is no a universal gestures language [9]. In addition, the highly articulated character, the human hand is deformable, and these troubles are in the issue of lacking good mathematical models for a top-down approach [10], [6], [8], [1] and references therein for meaningful cornerstones). In this paper, the attention is focused on the right positioning for grasping and handling in robotic applications, in the framework of a bottom-up approach.

An efficient design of grasping and handling tasks requires a robust identification of the right positioning [4]. These processes can be performed in terms of the (planar or volumetric) pose identification of the current posture or, in more advanced stages, from structure-from-motion algorithms [1]. The second approach for complex articulated objects is not easy. It requires the introduction of biomechanical constraints about a changing geometry and the allowed rigid motions. Hence, its computer implementation is considerably harder than the first one.
The *model-based* postures recognition and gestures tracking requires geometric and kinematic models for their unambiguous identification [6]. The high number of d.o.f. (27) for a 3D model of the human hand [10], and the need of a real-time computer implementation, suggest a feedback between bottom-up and top-down approaches. Following the *appearance-based* approach, we shall begin by extracting meaningful geometric data from images (segments and junctions) to generate easily updateable models in terms of grouping of points and closing segments in virtual planar polygonals to determine regions. In this way, we label clouds of extracted points and polygonal regions, which simplifies the tracking of mobile objects without need of matching homologue points. A related approach to ours can be read in [7], but the complexity of our images has forced us to simplify some more advanced aspects relative to the shape, its computer implementation and its symbolic representation by means of a virtual skeleton

II. METHODOLOGY

A. Brief Description of the Marcus Hand

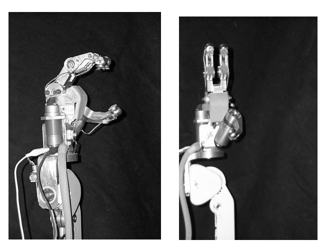

Figure 1 Two views of Marcus Hand

The Marcus robotic hand is assembled to a robotic arm in an anthropomorphic mobile robot, which is available at the Lab ARTS, Scuola Superiore Sant'Anna (Pisa, Italy. [3].

The Marcus hand was developed as a human prosthesis with three fingers and 2 degrees of freedom; actually, a mechanical coupling between two of the fingers allows them to be not completely dependent, so that the Marcus hand can be defined as having 2 ½ degrees of freedom ,. The thumb is equipped with an integrated fingertip comprising of a tactile array sensor, a thermal sensor and a dynamic sensor and it has only movement along the horizontal axis. The other fingertips are equipped with force sensor [2].

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25 Oct 2001	N/A	-

Title and Subtitle Monocameral Visual Recognition of Marcus Hand Postures for Personal Robotic Assistants	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Eca-Simm Group Department of Computer Science University of Valladolid, Spain	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group (UK) PSC 803 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom.	
Abstract	
Subject Terms	
Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU
Number of Pages 4	

B. Recognition Algorithm

The first version of our algorithm has been developed to work with a virtual or simulated robotic hand built with Open/GL [11] [5]. It is based on the extraction of meaningful geometric data concerning to the boundaries and discontinuities involving to the first derivatives. It is necessary to introduce geometric constraints for the grouping of mini-segments. These constraints allow us to update the information relative to the current state of phalanges and to patch together the geometric data along the fingers. Nevertheless, they will not be considered here, these algorithms are enough robust to allow us the postures recognition of a human hand [4].

Our recognition process of the hand has the following phases:

- *Low-level recognition:* acquisition of lists of points and selection of meaningful points.
- *Grouping* to *extract higher dimensional* characteristics (boundaries and regions)
- *Identification* of phalanges and fingers based on models
- *Symbolic Meaning* of this identification in terms of the relative pose of skeleton

In the low-level recognition phase, we extract contours and junctions as discontinuities for the intensity function in the grey scale. We have selected SUSAN filter [12] by its rapidity and efficiency. Susan filter has parameters to be adapted by the user, as the brightness threshold. It is useful when the ambient conditions change.

After applying the Susan Filter, we obtain a list of corners points. We extract only some good representatives of each small region by using a proximity and non-redundant threshold (based in empirical work) to determine that two points are neighbours. The model-based information allows us to discriminate between points located at boundaries and points located at the interior regions. Such discrimination is the key to connect boundary points in a consecutive way, but there are no universal procedures to close contours, not even in the piecewise linear case.

When the low-level recognition process is finished, the next step is the determination of the base of the hand (wrist) and the identification of each phalanx. We have implemented an algorithm that identifies each finger and their phalanx from the wrist of the hand. First, our identification algorithm creates a tree whose nodes represent a final point of a segment and whose links correspond to a segment, following natural adjacency criteria. From this tree, we identify the relative pose of each phalanx at the corresponding finger by using large segments obtained along the low-level recognition process Furthermore, the identification tree allows us to discriminate the wrist and the thumb. In this way, we obtain additional information about the relative position-orientation of all the system.

III. RESULTS

The visual devices and computer implementation are based on NETSIGHT, which is a vision system working on the Windows CE operating system. It is equipped with a Pentium MMX, grabber video card, software and some libraries of vision (MVT tools) already pre-implemented.

The Marcus hand can be moved in different ways according to its d.o.f. and movements of the arm supporting the hand. A static camera is located in front of the hand. To simplify the analysis, the images are taken from camera in a front-parallel view. However, our algorithms are enough robust to allow the posture identification for another more complicated relative orientations [4] [5] [11].

In a front parallel view, the camera can see two of the virtual two-in-one finger, one motor, and a tactile sensor in the thumb, a wrist that joins Marcus hand with robot arm and some cables. The Marcus hand is made of aluminium; material that reflects the light. There are some wires and motors, which make difficult the recognition of the contour and the base of the hand.

A. Recognition process

After the capture process, we have a set of coloured images, which we transform in pgm format. Then we have employed SUSAN filters [12] to extract corners and edges, by selecting thresholds for parameters involving the brightness and distance, and we have selected upper bounds for the maximal number of corners and edges.

Unfortunately, the information of segments is not enough meaningful for the identification tree, and it is necessary incorporate additional grouping criteria arising from a geographic analysis of data contained in the views.

B. Grouping Points

Instead working with the silhouette of the hand, we introduce some grouping criteria for clouds of points in terms of their gravity's centre or barycentre. In this way, we are able of working with a big number of small segments, becoming a serious trouble for segmentation and grouping. All these small elements appearing into the information processing are linked to the electromechanical structure[1]. These sets[2] of points are where there is some piece of union (screw, nut or motor) of the hand, because the processing of

[1] wires, sensors, motors, mechanical structure
[2] clouds

this zone gives rise to a great number of small segments; it causes these sets of corners.

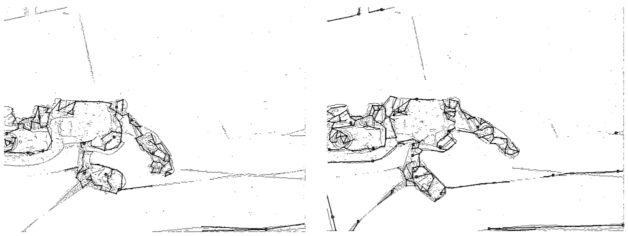

Figure 2 Two examples of grouping points

The camera does not move although the hand if that does it. This fact facilitates to maintain a window to us where to explore points in the image. The process begins by identifying the grouping point, which correspond to the hand wrist. For it, we have analysed those points of groupings that are in a middle of the left part of the image.

From this point, we can identify each part of the hand, deleting some additional grouping points that do not interest in that structure and looking for within the corners' list, those others that have not been identified. It is essential create a simple and realistic structure of the hand according with different parts that detect our algorithm. This allows us to recognize hand posture in a precise way.

In a front-parallel view, we detect two fingers: one in the upper position and the other one in the lowest position (thumb). Each finger has an anchor point with the base point (wrist) and includes at least various segments that correspond with carpals, metacarpals and phalanges. We incorporate two elements more to this structure because in the lower finger, Marcus hand has a motor and a union between the motor and first phalanx. In addition, we consider some movements of each part of hand structure that allows us in the correspondence phase. Once, we have identified all grouping points as hand structure elements, it is simple to know which posture have, just knowing which is the movement of the hand structure.

In this way, we have a fast and simple procedure based on the local topology of the view, which avoids to carry out long and expensive mathematical operations to find homologue points and to group segments by improving older results ([4]). In addition, it is very simple to describe the structure of any robotic hand. It suffices to identify and track each zone and its movement.

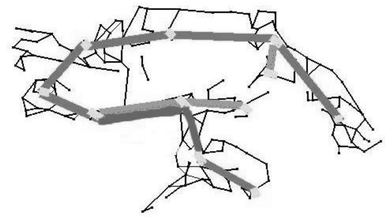

Figure 3: Grouping points conforms Hand Skeleton

IV. PERSONAL ROBOTIC ASSISTANTS

The described work allows to obtain a visual servoing for grasping and manipulation that can increase the performances of robots expected to operate in unstructured environments. This is the case, for example, of Personal Robotics, which is now the challenge of advanced robotics research worldwide [3]. Current frontiers of robotics research are in the development of anthropomorphic schemes for perception and behaviour and in the application of robots in daily life of common people. An initial step in this direction has been moved by Rehabilitation Robotics, which tried to apply robotic systems in the assistance of motor disable people since the late 80s. Even though first experiences were more replications of fixed robotic workstations in structured environments [2], further developments include mobile robotics and the need for operating in unstructured environments and therefore the need for robust perception-action closed loops.

V. DISCUSSION

Addressing the real-world case of recognising the posture of the Marcus hand, the work presented in this paper is based on the low-level appearance-based approach in the quasi-static case (allowing only very simple motions in front-parallel view), also due to the lack of enough flexible and easily updateable mobile data structures associated to the artificial Marcus hand.

Instead of identifying surfaces and volumetric primitives, we have implemented a skeleton-based approach to achieve a high-speed low cost implementation. In [4] two of the authors use marked fingertips to estimate their relative 3D position from stereo vision. Instead, we use simplified one-dimensional segment-based models that are generated in a symbolic way from a selective grouping of points and segments. In this way, we expect to improve the man-machine interaction in Personal Robotics, when robots are introduced in unstructured environments in presence of human persons, such that the assistance of severely disabled persons.

VI. CONCLUSIONS AND FUTURE WORK

The described process gains in rapidity and simplicity, because it is not necessary carry out long and expensive mathematical operations with a great number of points and segments. In addition, it is very simple to describe the structure of any robotic hand as we have done it. It is just necessary to detail each zone and its movement

The hand tracker will specify the pose of the projected hand in the image. By diminishing the error w.r.t. to an expected typical posture (a gesture in the next future), we simplify the tracking to the image level based on additional sensors relative to geometric (position-orientation) or dynamic (force, tactile) sensors. The kinematic properties of trajectories at the image are not easy to identify for complex articulated objects, such an anthropomorphic hand, e.g.. Instead, we are developing a system based on additional sensors that will allow us to compare the current geometric-dynamic configuration with the recognized one from the artificial viewpoint. The evaluation of errors would provide optimisation criteria and algorithms to perform a better design of supervised learning. Our next goal is to incorporate biomechanical constraints into a supervised learning to improve the man-machine interaction.

A far-reaching goal is related to the 3D reconstruction of the current pose, and the motion tracking for some selected points or segments located at thumb and another finger. The real-time identification would be applied to a simulated robot gripper, before introducing in biomedical applications. Currently, we have obtained meaningful results about the identification and tracking based on global kinematic characteristics of regions, to overimpose finer geometric and dynamic information about points and segments.

Acknowledgments

This work has obtained a financial support from a European TMR Grant within the MOBINET Project, carried at the ARTS (Advanced Robotics Technology and Systems) Lab of the Scuola Superiore di Sant'Anna, Pisa, Italy, and from the Acción-Integrada Hispano-Italiana 2000/0090.

References

[1] N.Andreff, R.Horaud and B.Espiau: "Robot-Hand Eye calibration using Structure-from-Motion", INRIA TR Febr.2000.

[2] P. Dario, C. Laschi, E. Guglielmelli, "Sensors and actuators for ' humanoid' robots", RSJ International Journal Advanced Robotics, Special Issue on Humanoid, Vol.11, No.6, 1998, VSP Publishing, The Netherlands, pp.567-584.

[3] P. Dario, E. Guglielmelli, C. Laschi, G. Teti, "MOVAID: a personal robot in everyday life of disabled and elderly people", Technology and Disability Journal, No.10, 1999, IOS Press, The Netherlands, pp.77-93.

[4] J.Finat and M.Gonzalo-Tasis,"Fast recognition of postures for a simplified three-fingered artificial hand", Proc. IEEE SMC San Diego, October 1998.

[5] M. Gonzalo-Tasis. D. Sánchez, R. Pellón, and J. Finat , "Symbolic Models For Postures Recognition Of A Three Fingered Artificial Hand", ISPRS 2000, Amsterdam, 16-23 July 2000

[6] J.Lee and T.Kunii: "Model-Based Analysis of Hand Postures", IEEE Computer Graphics and Applications, Sept 1995, 77-86

[7] J.Martin and J:L:Crowley: "An appearance-based approach to Gesture Recognition", 9th Intl Conf on Image Analysis and Processing (Florence, Italy, Sept, 1997).

[8] H.Ouhaddi and P.Horain: "Design, adjustement and registration of a 3D articulated model for hand tracking in image sequences", Technical Report, 1998.

[9] V.I.Pavlovic, R.Sharma and T.S.Huang: "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review", IEEE Trans. PAMI 19(7) July 1997.

[10] J.M.Rehg and T.Kanade: "Digit eyes: vision-based human hand tracking", Technical Report CMU-CS 93-220. Carnegie Mellon University, 1993.

[11] D. Sánchez, R. Pellón, "Reconocimiento de posturas de una mano artificial", Master Thesis, Dept. of Computer Science, University of Valladolid, February 1999

[12] S.M. Smith and J.M. Brady. "SUSAN - a new approach to low-level image processing". *Int. Journal of Computer Vision*, 23(1):45--78, May 1997.